Promptify: Text-to-Image Generation through Interactive Prompt Exploration with Large Language Models

Abstract

Text-to-image generative models have demonstrated remarkable capabilities in generating high-quality images based on textual prompts. However, crafting prompts that accurately capture the user's creative intent remains challenging. It often involves laborious trial-and-error procedures to ensure that the model interprets the prompts in alignment with the user's intention. To address these challenges, we present Promptify, an interactive system that supports prompt exploration and refinement for text-to-image generative models. Promptify utilizes a suggestion engine powered by large language models to help users quickly explore and craft diverse prompts. Our interface allows users to organize the generated images flexibly, and based on their preferences, Promptify suggests potential changes to the original prompt. This feedback loop enables users to iteratively refine their prompts and enhance desired features while avoiding unwanted ones. Our user study shows that Promptify effectively facilitates the text-to-image workflow, allowing users to create visually appealing images on their first attempt while requiring significantly less cognitive load than a widely-used baseline tool.